



Máster Universitario en Sistemas Inteligentes

TRABAJO DE FINAL DE MÁSTER

Detección de comienzos de notas musicales en archivos de audio: estudio experimental

Autor:
Salvador Martí Solsona

Tutor académico:
Ramón Mollineda Cárdenas

Fecha de lectura: 26 de Noviembre de 2015
Curso académico 2014/2015

Resumen

El desarrollo de soluciones computacionales aplicadas a problemas relacionados con el sonido y la música es un área de investigación emergente.

Actualmente existen múltiples ámbitos de problemas en esta área del conocimiento, entre ellos: el procesamiento de audio, desarrollo de interfaces musicales, modelización de la interpretación musical, descripción del sonido y de la música, etc. Entre las disciplinas científicas más populares para tratar con estos problemas podemos encontrar el Aprendizaje Automático, Tecnologías Semánticas, Interacción Hombre Máquina y Procesado de Señal.

Un problema actual es el llamado genéricamente '*Onset detection*' o detección en sonido de inicios de nota. El objetivo final de todo el proceso es conseguir una lista de instantes de tiempos en los que, presumiblemente, comienzan las notas. Existen multitud de aplicaciones que demandan una solución a este problema, como por ejemplo: la transcripción automática, el análisis de la estructura y frases musicales, la evaluación de la similitud musical, el análisis del rendimiento expresivo y la recuperación de información musical.

Como parte de un Trabajo de Introducción a la Investigación previo al desarrollo de este proyecto, se ha revisado el estado del arte más reciente del problema '*Onset detection*' y se ha presentado un conjunto de datos en un formato muy cercano al contexto real que permitirá ilustrar este problema. También se ha planteado un proceso de re-ordenación de estos datos a un formato más estándar que favorezca su análisis y se ha descrito la metodología de evaluación de la efectividad de los métodos que se usará. Estas tareas son, a grandes rasgos, las que han permitido obtener una base de conocimiento sobre el problema y sus soluciones potenciales, y acometer el estudio experimental que propone este Trabajo Final de Máster.

Este proyecto pretende diseñar un marco experimental que permita evaluar soluciones al problema '*Onset detection*'. Cada solución contemplará, potencialmente, métodos de transformación de los datos, de aprendizaje y detección (clasificación), y de análisis de la efectividad de los resultados. Finalmente, se realizará una comparación de los métodos involucrados, con el fin de proponer directrices que ayuden en futuras investigaciones relacionadas.

Palabras Clave

Música, detección automática, inicio de nota, aprendizaje automático, procesamiento de señal

Keywords

Music, automatic detection, onset, machine learning, signal processing

Índice general

Capítulo 1: Introducción	5
1.1. Motivación	5
1.2. Resumen del estado del arte	6
1.3. Objetivos	7
1.4. Organización del documento	7
Capítulo 2: Definición y caracterización del problema	9
2.1. Definición de onset	9
2.2. Anotaciones manuales	10
2.3. Planteamiento del problema de detección automática del onset	10
2.4. Naturaleza de la clase y desbalance	12
Capítulo 3: Metodología	13
3.1. Tratamiento automático de la asimetría	13
3.2. Preparación de conjuntos de entrenamiento: Leaving one FILE out	14
3.3. Clasificadores	14
3.4. Estimación fiable de medidas de rendimiento	16
Capítulo 4: Análisis de resultados	19
4.1. Resultados individuales	19
4.2. Resultados generales	30
Capítulo 5: Conclusiones y trabajo futuro	33
Referencias	35

Capítulo 1: Introducción

En este primer capítulo se describen los motivos por los que se consideró oportuno investigar la detección automática de comienzos de notas musicales en ficheros de audio. Además, se introduce brevemente el estado del arte del problema, se presentan de manera concisa los objetivos principales del trabajo y finalmente se detalla como se ha organizado el presente documento.

1.1. Motivación

La música se define como el arte de organizar sensible y lógicamente una combinación coherente de sonidos y silencios utilizando los principios fundamentales de la melodía, la armonía y el ritmo, mediante la intervención de complejos procesos psicoanímicos. El objetivo de este arte es provocar una experiencia estética en el oyente, y expresar sentimientos, emociones, pensamientos o ideas. La música es, por tanto, un estímulo que afecta a la percepción del individuo. El oyente percibe, interpreta, siente y responde a cada estímulo musical. Ejemplos claros y evidentes son gestos como mover la cabeza, llevar el ritmo con los pies o chasquear los dedos.

En todo caso, la atención del individuo se centra en cada nota de la secuencia musical, la cual ocurre tras un cambio o transición desde la nota que le precede. Sin cambio, no puede haber significado musical. Por lo tanto, una condición necesaria que permite al ser humano comportarse de esta forma es su capacidad para percibir los puntos de partida de las notas musicales.

Existen estudios computacionales relacionadas con el sonido y la música que tratan de imitar, de una manera artificial, los mecanismos naturales del ser humano para descubrir estos inicios de notas musicales. Entre las disciplinas científicas más populares que abordan esta cuestión podemos encontrar el Aprendizaje Automático, Tecnologías Semánticas, Interacción Hombre-Máquina y el Procesado de Señal.

Una de las denominaciones que recibe este problema en el ámbito del procesamiento de señales es '*Onset detection*', el cual es cada vez más estudiado por parte de la comunidad científica. Formalmente, esta tarea se define como la detección automática del instante de inicio de un evento en una señal. Un subdominio muy tratado en la literatura es el llamado '*Audio Onset Detection*' o '*Music Onset Detection*', el cual se centra en la detección automática de inicios de nota en la señal de audio. El objetivo de este proceso es conseguir una lista de instantes de tiempos en los que, presumiblemente, comienzan las notas.

Para concluir, la detección automática de inicios de nota en señales de audio abre nuevas posibilidades a multitud de aplicaciones que demandan una solución a este problema como, por ejemplo, la transcripción automática, el análisis de la estructura de frases musicales, la evaluación de la similitud musical entre dos o más piezas, el análisis del rendimiento expresivo o la recuperación de información musical. Por estos motivos se considera de gran interés el estudio de este problema y la realización de trabajos como el que recoge este documento.

1.2. Resumen del estado del arte

Actualmente existen dos planteamientos distintos que abordan la detección de inicios de nota en señales de audio: el primero propone utilizar únicamente técnicas de procesamiento de señal mientras que el segundo incluye trabajos que aplican métodos de aprendizaje automático.

Las propuestas que usan puramente técnicas de procesamiento de señal tratan de encontrar la característica o características físicas, o ecuaciones, que mejor describen el evento. Investigaciones que se han centrado en este tipo de técnicas han planteado varias características a tener en cuenta como, por ejemplo, la detección de cambios en el dominio de la energía y de la fase de las señales sonoras [1]. El estudio de los cambios de energía es más útil en la detección de comienzos bien definidos, generalmente producidos por instrumentos de percusión, mientras que los cambios basados en fase son capaces de detectar inicios suaves, que denotan el comienzo de una excitación pero pueden no indicar un cambio en la energía de la señal. En otra investigación [2] se ha propuesto el uso combinado de la información de la energía y de la fase para la detección de comienzos de notas musicales. Este método mejora los resultados de los enfoques separados basados en energía o fases.

Por otra parte, las técnicas de aprendizaje automático se utilizan para construir las mejores fronteras de decisión entre fragmentos de audio que contienen inicios de nota y fragmentos que no los contienen. Estas técnicas son aplicadas comúnmente sobre características de audio para detectar los patrones de inicios de notas en un dominio de dimensiones superiores. Por poner algún ejemplo, en [3] se investiga las capacidades de dos métodos de aprendizaje automático como son las máquinas de vectores soporte (SVM) y las redes neuronales (NN) en función de dos parámetros del contexto musical (tono y *tempo*). Además, en ese estudio se proponen distintas formas de incorporar información del tono y del *tempo* a las técnicas de aprendizaje automático. Otros estudios sugieren diferentes métodos como, por ejemplo, el uso de los *k* vecinos más próximos [4] o técnicas basadas en *boosting* [5].

Todos estos métodos presentados han sido evaluados y probados en entornos fuera de línea (*offline*). Una tendencia en esta área de investigación, que ha tomado bastante importancia en los últimos años, ha sido la de modificar los métodos ya existentes para permitir una ejecución en línea (*online*). Estos estudios han concluido que los métodos de detección basados en la energía y la fase son adecuados para entornos *online* ya que son fáciles de implementar y presentan un coste computacional modesto. Por contra, los métodos basados en técnicas de aprendizaje automático dependen de un gran conjunto de datos de entrenamiento y demandan mayor uso de recursos computacionales con lo que, en su mayoría, son inadecuados para entornos *online*. Por poner un ejemplo de este tipo de técnicas, en [6] se comparan cuatro métodos de detección de comienzo de nota basadas en la energía de la señal y se propone un método para escenarios en línea.

1.3. Objetivos

El objetivo general de este proyecto es la evaluación de métodos de aprendizaje automático en la solución del problema '*Onset detection*'. Los objetivos concretos son los siguientes:

- Evaluar técnicas de pre-procesado de datos atendiendo a la efectividad para gestionar las complejidades subyacentes.
- Evaluar técnicas de aprendizaje y clasificación atendiendo a la efectividad para resolver el problema de interés.
- Identificar herramientas adecuadas para la visualización y comparación de resultados.
- Proponer directrices que permitan abordar de forma integral y efectiva el problema de '*Onset detection*'.

1.4. Organización del documento

En esta sección se describen los distintos capítulos que componen la estructura de la memoria:

- **Definición y caracterización del problema:** se define el concepto de *onset* y se explica cómo plantear un problema genérico de detección automática de *onsets*. Para acabar, se discuten ciertas complejidades inherentes al problema.
- **Metodología:** se describe la metodología seguida para preparar y generar el marco experimental necesario para llevar a cabo este trabajo.
- **Análisis de resultados:** se realiza una recopilación y análisis de los resultados obtenidos.
- **Conclusiones y trabajo futuro:** se discuten las conclusiones finales después de la investigación realizada y se plantean directrices de trabajo futuro en el tema del proyecto.

Capítulo 2: Definición y caracterización del problema

En este segundo capítulo se resume los conceptos teóricos básicos del problema, que se abordaron en el anterior Trabajo de Iniciación a la Investigación (TII). En primer lugar, se define el concepto de *onset*. A continuación, se sintetiza cómo plantear un problema genérico de detección automática de *onsets* a partir de una colección de datos con una organización muy cercana a la realidad del fenómeno musical. Finalmente, se presentan ciertas complejidades inherentes al problema.

2.1. Definición de *onset*

Como se ha comentado en el capítulo anterior, la localización y detección de *onsets* es un procedimiento muy útil para un gran número de aplicaciones.

Un *onset* se define como el momento en el que comienza una nota musical dentro de una señal de audio. La idea de *onset* es bastante ambigua por lo que cada investigador suele hacer una interpretación diferente del instante de tiempo que se debe etiquetar. Es buen momento para introducir y hacer una clara distinción entre tres conceptos físicos relacionados y que pueden llevar a confusión: ataque, transitorio y *onset*. La Figura 1 muestra cómo se podría diferenciar estos conceptos en el caso simple de una nota aislada.

- Ataque: Para una nota es el intervalo de tiempo en el que la envolvente de amplitud incrementa.
- Transitorio: Este concepto es más complejo de describir con precisión. Informalmente, los transitorios son intervalos cortos durante los cuales la señal evoluciona rápidamente, en muchas ocasiones de una forma impredecible y no trivial. Para el caso de los instrumentos musicales, el transitorio a menudo coincide con el periodo durante el cual la excitación (por ejemplo, un golpe de martillo en un piano) tiene lugar y luego se amortigua, dejando sólo la lenta decadencia en las frecuencias de resonancia del cuerpo. Un problema importante es el tiempo de resolución útil, ya que el oído humano no puede distinguir entre dos transitorios separados menos de 10 ms.
- Onset: Es un único instante de tiempo discreto elegido para marcar el origen de un transitorio. En muchos casos coincide con el inicio de un transitorio, o con el instante de tiempo más temprano en el que el transitorio puede ser detectado de forma fiable.

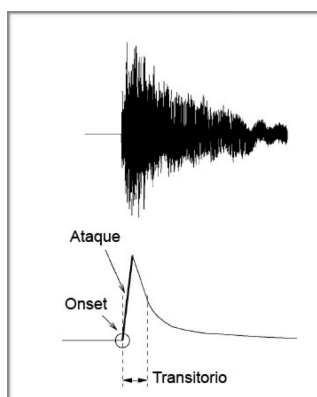


Figura 1: Ataque, transitorio y *onset* en el caso ideal de una nota simple.

2.2. Anotaciones manuales

Para poder implementar cualquier sistema de detección automática se necesita, además de las señales a analizar, los instantes de tiempo anotados por personas expertas en los que ocurren los *onsets* en cada señal. Estos tiempos de referencia se obtienen mediante un proceso manual y laborioso, que requiere tiempo y concentración por parte del experto que lo realiza.

El etiquetado manual de *onsets* en señales de audio se refiere al establecimiento de los instantes exactos de tiempo en los que un oyente humano considera que existen comienzos de notas, también denominado *ground truth*. En la literatura se puede observar que las mismas señales de prueba, etiquetadas por diferentes personas, producen instantes de referencia diferentes, mostrando una gran dependencia del método utilizado para etiquetar, de las características de la señal, y del propio oyente. De estos estudios se concluye que la evaluación de un sistema de detección automática de *onsets* en señales de audio es una cuestión compleja y, por lo tanto, nada trivial.

2.3. Planteamiento del problema de detección automática del *onset*

El objetivo de esta sección es explicar cómo plantear un problema genérico de detección automática de *onsets* (abstracción), a partir de una colección de datos con una organización muy cercana a la realidad del fenómeno musical.

Cada fichero de audio se divide en pequeños segmentos (*frames*) contiguos de 46 milisegundos de duración, que posteriormente serán analizados para comprobar si contienen o no un inicio de nota (detección de *onset*).

Desde un punto de vista del reconocimiento de patrones clásico, cada segmento será tratado como un prototipo descrito por un vector de diez características que representan distintos aspectos físicos de las señales de audio y, usando la información del *ground truth*, cada *frame* será etiquetado como *onset* (contiene un inicio de nota) o no *onset* (no contiene un inicio de nota). De esta forma, la tarea original se habrá convertido en un problema equivalente de dos clases. El aspecto o formato al cual llegamos es típico en aprendizaje automático y queda reflejado en la Figura 2.

Las etiquetas de los *frames* vienen determinadas por los instantes de tiempo del *ground truth*. Como ya se ha comentado previamente, estos aparecen en ficheros de texto aparte y tienen el mismo nombre que el fichero de audio correspondiente. Por cada línea existe una notación temporal que indica el instante en el que comienza una nota (*onset*) en el fichero de audio relacionado. Por lo tanto, la cantidad de líneas representa la cantidad de *onsets*. El problema es que estos datos no están alineados con la resolución de segmentos establecida (46 ms), como consecuencia de que son anotaciones totalmente manuales. Mediante varios *scripts*, desarrollados en el pasado TII, se estableció una correspondencia entre estas anotaciones y los *frames* extraídos del fichero de audio, atendiendo a varios criterios de clasificación. El primer criterio, y el más cercano a la definición de *onset*, consiste en alinear cada tiempo anotado con el *frame* más próximo, según su posición temporal en el fichero de audio. A este criterio se le ha llamado etiquetado estándar. El segundo criterio, que supone asumir cierto error o incertidumbre en las anotaciones manuales, marca como *onset* a todos aquellos *frames* cuyos instantes

de tiempo disten a ± 50 ms de la notación temporal del *onset* más próxima. A este criterio se le ha llamado etiquetado extendido.

La colección consta de quince ficheros de audio, descritos en la Tabla 1, con sus respectivos archivos de *ground truth*.





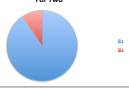
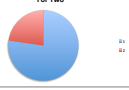
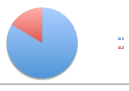

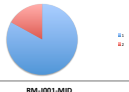







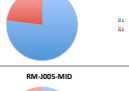
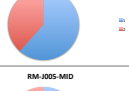

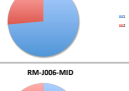
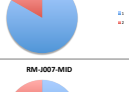
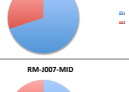

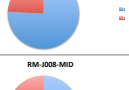
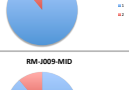
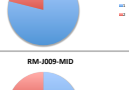
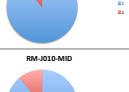
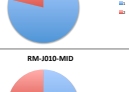


Nº	Título	Nº total de frames	Nº Onsets Et. Estandar	Proporción Onsets Et. Estandar	Nº Onsets Et. Extendido	Proporción Onsets Et. Extendido
1	Abyss	4064	688		1734	
2	Crescent Serenade	3608	992		1863	
3	For Two	8096	843		2383	
4	Jive	4368	856		1770	
5	Lounge Away	3424	700		1387	
6	RM-J001-MID	4376	859		1734	
7	RM-J002-MID	8128	867		1807	
8	RM-J003-MID	3520	691		1349	
9	RM-J004-MID	3584	1069		2234	
10	RM-J005-MID	3984	709		1436	
11	RM-J006-MID	3432	690		1476	
12	RM-J007-MID	4672	818		1494	
13	RM-J008-MID	4744	590		1266	
14	RM-J009-MID	4760	595		1283	
15	RM-J010-MID	4528	571		1229	

Tabla 1: Piezas musicales analizadas. En rojo la proporción de muestras etiquetadas como *onset* con respecto al total de muestras.

muestra 1	Atributo 1	Atributo 2	...	Atributo 10	Etiqueta de clase
muestra 2	Atributo 1	Atributo 2	...	Atributo 10	Etiqueta de clase

Figura 2: Formato final para el correcto procesado de un sistema de aprendizaje automático.

2.4. Naturaleza de la clase y desbalance

Como ya se ha comentado en anteriores secciones, es posible dividir una señal de audio en muestras (*frames*) de unos pocos milisegundos de duración y a partir de aquí clasificar cada una de ellas en función de si contiene un *onset* o no.

Una característica muy importante es que se esperan muchos más *frames* que no tienen *onsets* que *frames* que sí los contienen. Por lo tanto, estamos ante un problema de clasificación desbalanceado. Un primer motivo es la propia naturaleza y definición de *onset*, ya que especifica un único instante de tiempo discreto elegido para marcar el origen de un transitorio. Un segundo motivo es el método automático elegido para alinear los tiempos de *ground truth* con los tiempos de las muestras. Si el criterio es hacer corresponder cada anotación de tiempo del *ground truth* con una única muestra en la señal de audio original, se conseguirá un etiquetado más próximo a la propia definición de *onset* y a la propia intuición del oyente experto (etiquetado estándar), aunque se obtendrá una cantidad minoritaria de muestras etiquetadas como *onset*.

Sin embargo, también sería posible asumir cierta incertidumbre en los tiempos seleccionados por el observador experto, ya que la elección de los instantes de tiempo del *ground truth* depende del oyente y de las técnicas concretas que hayan utilizado. En este caso, dado un instante de *ground truth*, además del *frame* más próximo, podrían etiquetarse como *onset* aquellos *frames* contiguos suficientemente próximos al instante de *ground truth* según un umbral. Esta forma de etiquetar se ha denominado en este trabajo etiquetado extendido.

En el pasado TII se realizó un estudio cuyos resultados demuestran que existe un desbalance importante para el problema original. Cada audio presenta sus propias características musicales que lo hacen único y que influyen en el número de inicios de notas finales detectados. Se puede resumir que el desbalance medio para el conjunto de ficheros ronda la relación 1:5 en el caso del etiquetado estándar (aproximadamente el 17% de *onsets*). Por otro lado, se observa que con el etiquetado extendido se obtiene una relación media de 1:1,7 (aproximadamente el 37% de *onsets*).

Capítulo 3: Metodología

En este tercer capítulo se describe la metodología seguida para preparar y generar el marco experimental necesario para llevar a cabo este trabajo. El primer paso en este proceso ha sido el tratamiento automático de la asimetría sobre los ficheros de audio mediante diferentes técnicas ya presentadas en el pasado TII y la generación de nuevos conjuntos de entrenamiento con clases balanceadas. En segundo lugar se ha modelado una serie de escenarios de experimentación que permitirán establecer diversas pruebas sobre cada fichero. Seguidamente, se ha realizado un proceso de detección automática de *onsets* (clasificación binaria) sobre estos escenarios, cuyos resultados han dado lugar a ciertas métricas que valoran el rendimiento final del sistema.

3.1. Tratamiento automático de la asimetría

La asimetría entre clases puede ser una dificultad añadida para los detectores automáticos de *onsets* a la hora de generar un modelo de clasificación lo más preciso posible. Incrementando de manera artificial muestras de la clase minoritaria (SMOTE) [7] o reduciendo muestras de la clase mayoritaria (*Undersampling*) [8], se puede ayudar a aumentar la robustez del modelo y, consecuentemente, a mejorar el rendimiento del sistema. Por este motivo, se realizó un proceso previo dónde se ha ampliado la base de datos de partida con nuevos ficheros de entrenamiento, producto de aplicar estas técnicas de re-muestreo sobre los conjuntos de datos originales de los archivos de audio con etiquetado estándar y extendido.

Etiquetado		Método de desbalance
Fichero	Estándar	Neutro
		SMOTE
		<i>Undersampling</i>
	Extendido	Neutro
		SMOTE
		<i>Undersampling</i>

Figura 3: Técnicas de tratamiento de la asimetría utilizadas para cada combinación de archivo de audio y tipo de etiquetado.

Como se puede apreciar en la Figura 3 se definen seis tipos de ficheros por cada combinación de archivo de audio y técnica de etiquetado. Por otro lado, el término ‘neutro’ hace referencia a la no aplicación de ninguna técnica de re-muestreo, es decir, al uso del conjunto de datos original en cada caso. Por último, es interesante indicar que para los nuevos ficheros basados en SMOTE y *Undersampling* se han creado múltiples versiones. La razón se explicará más adelante en la sección 3.4.

3.2. Preparación de conjuntos de entrenamiento: *Leaving one FILE out*

Para preparar los conjuntos de entrenamiento se ha adaptado una técnica muy utilizada en el ámbito de la inteligencia artificial y que recibe genéricamente el nombre de '*Leaving one out*'. Esta técnica consiste en separar los datos de forma que para cada iteración se tenga una sola muestra de prueba. Sin embargo, en nuestro caso, en lugar de una muestra, se apartará un archivo de audio íntegro, mientras el resto de archivos conformará el conjunto de datos de entrenamiento. Esta es la razón por la que hemos propuesto el nombre '*Leaving one FILE out*'. Por lo tanto, y de manera generalizada, cada fichero original se considerará conjunto de *test* en múltiples experimentos en los que el resto de ficheros de audio, a través de sus versiones resultantes de un mismo tipo de remuestreo, se agruparán para formar conjuntos de entrenamiento. En la Figura 4 se puede ver el esquema general para el problema.

La primera columna representa todos los ficheros originales con etiquetado estándar. La segunda columna denota los nuevos ficheros generados para cada archivo de audio original según una combinación concreta de tipo de etiquetado y método de desbalance (por ejemplo, etiquetado estándar + SMOTE). Por otro lado, la tercera columna representa al archivo original desempeñando el rol de conjunto de prueba en cada iteración y, finalmente, el último grupo de columnas muestra la combinación de las versiones remuestreadas de ficheros que se unen para formar el conjunto de entrenamiento complementario al archivo original elegido como *test*.

3.3. Clasificadores

En esta sección se describen brevemente los algoritmos de clasificación utilizados para llevar a cabo la detección automática de *onsets*. Los algoritmos han sido básicamente dos: '*K-nearest neighbour*' (K-NN) [9] para $K=1,3$ y '*Support Vector Machine*' (SVM) [10].

El algoritmo K-NN va a clasificar un nuevo *frame* o muestra en la clase más frecuente entre sus K vecinos más próximos dentro del conjunto de entrenamiento. Se trata de una idea muy sencilla, intuitiva y fácil de implementar, por lo que es un algoritmo muy extendido. Además, K-NN se define como un algoritmo perezoso (*lazy*) ya que durante el entrenamiento sólo guarda las instancias de manera local y no construye ningún modelo global. En cuanto a sus limitaciones, es muy sensible a los atributos irrelevantes, al incremento de la dimensionalidad, al ruido y su rendimiento depende de la función de distancia que implemente.

K-NN es, en cierta medida, complementario a las SVM ya que estas últimas son capaces de construir un modelo global a partir de ejemplos de entrenamiento y predecir la clase de las nuevas muestras. Dicho modelo consiste en un hiperplano que separe de forma óptima las muestras de dos clases, potencialmente proyectadas sobre un espacio de mayor dimensionalidad. El criterio de optimalidad está determinado por la búsqueda del hiperplano que maximiza la distancia (margen) a las muestras más próximas de ambas clases, las cuales se denominan vectores soporte.

Ficheros originales		Etiquetado + remuestreo		Folds															
Etiquetado estándar		Test		Training															
Crescent Serenade	Abyss	Abyss*	Abyss	Crescent Serenade*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*		
	FortTwo	Crescent Serenade*	Crescent Serenade	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*		
	Jive	FortTwo*	FortTwo	Abyss*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*			
	LoungeAway	LoungeAway*	LoungeAway	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*		
RM-1001-MID		RM-1001-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1002-MID		RM-1002-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1003-MID		RM-1003-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1004-MID		RM-1004-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1005-MID		RM-1005-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1006-MID		RM-1006-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1007-MID		RM-1007-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1008-MID		RM-1008-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1009-MID		RM-1009-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	
RM-1010-MID		RM-1010-MID*		Crescent Serenade*	Abyss*	FortTwo*	Jive*	LoungeAway*	RM-1001-MID*	RM-1002-MID*	RM-1003-MID*	RM-1004-MID*	RM-1005-MID*	RM-1006-MID*	RM-1007-MID*	RM-1008-MID*	RM-1009-MID*	RM-1010-MID*	

Figura 4: Esquema general del ‘Leaving one FILE out’.

3.4. Estimación fiable de medidas de rendimiento

Un experimento va a consistir en entrenar un cierto clasificador mediante un conjunto de entrenamiento como el descrito en la Sección 3.2, y la clasificación del conjunto de pruebas (o *test*) correspondiente (un archivo de audio original). Como resultado del proceso, se obtendrá una serie de predicciones de etiquetas de clase para las muestras de *test*. A partir de aquí se inicia un proceso que consiste en comparar, muestra a muestra, las etiquetas de clase reales asociadas originalmente a las muestras de test con las predicciones realizadas por el clasificador. De esta forma se calculan los índices *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) y *False Negative* (FN). A partir de aquí es posible calcular de manera sencilla los valores para las medidas $Recall = TP / (TP + FN)$, $Precisión = TP / (TP + FP)$, $Especificidad = TN / (TN + FP)$, $F-measure = (Recall * Precisión) / (Precisión + Recall)$ y $Accuracy = (TP + TN) / (TP + FP + TN + FN)$, cuyos significados aparecen descritos en la memoria del TII previo a este TFM. La Tabla 2 resume los índices y medidas obtenidas por cada experimento.

TP	TN	FP	FN	Recall	Precisión	Especificidad	F-measure	Accuracy
----	----	----	----	--------	-----------	---------------	-----------	----------

Tabla 2: Índices y medidas obtenidas por cada experimento.

Teniendo en cuenta la componente aleatoria de los métodos de re-muestreo, cada tipo de experimento según configuraciones de datos definidas en la Sección 3.2, determinado por cierta combinación de un método de etiquetado y una estrategia de re-muestreo descrita en la sección 3.1, se repite 10 veces. El objetivo es calcular valores medios de las medidas de rendimiento explicadas antes, los cuales se suponen más fiables que los resultados individuales de experimentos particulares. La Figura 5 resume la cantidad de experimentos realizados:

TEST	Nº pruebas	Fichero de entrenamiento	
		Tipo de etiquetado	Método de desbalance
Fichero de audio original	10	Estándar	SMOTE
	10	Estándar	<i>Undersampling</i>
	1	Estándar	Neutro
	10	Extendido	SMOTE
	10	Extendido	<i>Undersampling</i>
	1	Extendido	Neutro

Figura 5: Número de pruebas realizadas por cada fichero de audio original.

A continuación se explica más detalladamente el significado de las repeticiones realizadas para clasificar un conjunto concreto de prueba, vinculado a un fichero de audio original:

- “**Etiquetado estándar + SMOTE**”, 10 repeticiones con 10 conjuntos de entrenamiento, cada uno de ellos creados a partir del uso de *SMOTE* sobre los 14 ficheros restantes con etiquetado *estándar*.
- “**Etiquetado extendido + SMOTE**”, 10 repeticiones con 10 conjuntos de entrenamiento, cada uno de ellos creados a partir del uso de *SMOTE* sobre los 14 ficheros restantes con etiquetado *extendido*.
- “**Etiquetado estándar + Undersampling**”, 10 repeticiones con 10 conjuntos de entrenamiento, cada uno de ellos creados a partir del uso de *Undersampling* sobre los 14 ficheros restantes con etiquetado *estándar*.
- “**Etiquetado extendido + Undersampling**”, 10 repeticiones con 10 conjuntos de entrenamiento, cada uno de ellos creados a partir del uso de *Undersampling* sobre los 14 ficheros restantes con etiquetado *extendido*.
- “**Etiquetado estándar + Neutro**”, 1 prueba con el conjunto de entrenamiento definido por la unión de las 14 versiones con el etiquetado *estándar* del resto de ficheros (excepto el mismo).
- “**Etiquetado extendido + Neutro**”, 1 prueba con el conjunto de entrenamiento definido por la unión de las 14 versiones con etiquetado *extendido* del resto de ficheros (excepto el mismo).

El motivo por el cual se decide realizar múltiples pruebas sobre las técnicas de *SMOTE* y *Undersampling* es por el hecho de que estos métodos presentan cierta aleatoriedad a la hora de generar o destruir muestras, por lo que se busca, al final del proceso, un resultado promedio menos dependiente de las singularidades de cada proceso aleatorio.

Finalmente, las medidas de rendimiento obtenidas para cada experimento se han analizado, en primer lugar, a nivel individual de fichero y, finalmente, a nivel general. Este análisis se presenta, desde estos dos puntos de vista, en el siguiente capítulo.

Capítulo 4: Análisis de resultados

En este cuarto capítulo se realiza una recopilación y análisis de los resultados obtenidos, los cuales han sido distribuidos en dos secciones. En la primera de ellas se presentan las soluciones individuales por cada fichero. En la segunda sección se realiza un resumen del rendimiento del proceso de detección automática de *onsets* a nivel general.

4.1. Resultados individuales

Para organizar los distintos resultados de clasificación que se obtienen por cada fichero de audio se utiliza una tabla como la que se puede visualizar en la Figura 6. En ella se pueden observar tres grandes bloques horizontales, uno por cada algoritmo de clasificación utilizado. De la misma manera, si se analiza la tabla de manera vertical se pueden apreciar tres grandes bloques, que corresponden con las tres estrategias de re-muestreo utilizadas, SMOTE, *Undersampling* y Neutro, respectivamente. Además, cada bloque vertical se vuelve a dividir en dos, en función de la técnica de etiquetado (estándar o extendido) que ha intervenido en la formación de los sucesivos archivos de entrenamiento. Se estructura, de esta forma, todo el conjunto de configuraciones de datos definidas en el anterior Capítulo 3 Sección 2. De esta organización surgen dieciocho unidades o bloques básicos por cada combinación de <algoritmo de clasificación - técnica de etiquetado - estrategia de re-muestreo>, constando cada una de ellas de una fila, de forma que:

- Media: Corresponde con el promedio de las diez repeticiones del experimento determinado por la combinación de métodos que define cada bloque.
- Prueba: Corresponde con el único experimento en el que interviene el conjunto de entrenamiento construido a partir de los ficheros originales, a los que no se les ha aplicado ninguna técnica de re-muestreo.

Finalmente, cada unidad o bloque se divide en nueve columnas que van a contener los valores de los índices y medidas de rendimiento presentadas en el Capítulo 3 Sección 4. A continuación, en la Tabla 3 se muestran los resultados individuales por cada fichero. Para su mejor visualización se ha dividido la estructura en tres partes, cada una de las cuales contiene los resultados obtenidos por estrategia de re-muestreo.

FICHERO	SMOTE						UNDER						NEUTRO					
	ESTANDAR			EXTENDIDO			ESTANDAR			EXTENDIDO			ESTANDAR			EXTENDIDO		
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA																		
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA																		
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA																		

Figura 6: Esquema general para la organización de los resultados de detección automática de *onsets* de un fichero particular.

FICHERO																				
SMOTE	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	296,1	2649	726,8	391,9	0,43	0,29	0,785	0,346	0,725		371,7	2157	1219	316,3	0,54	0,234	0,639	0,326	0,622	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	343,4	2491	885,2	344,6	0,499	0,28	0,738	0,358	0,697		388,9	2052	1324	299,1	0,565	0,227	0,608	0,324	0,601	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	470,6	3065	311,5	217,4	0,684	0,602	0,908	0,64	0,87		563	2701	675,3	125	0,818	0,455	0,8	0,585	0,803	

ABYSS																				
UNDER	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	374	2247	1129	314	0,544	0,249	0,666	0,341	0,645		392,7	1982	1394	295,3	0,571	0,22	0,587	0,317	0,584	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	358,4	2379	997,4	329,6	0,521	0,264	0,705	0,351	0,673		370,9	2063	1313	317,1	0,539	0,22	0,611	0,313	0,595	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	487,4	2994	382,3	200,6	0,708	0,561	0,887	0,626	0,857		546,7	2797	579,4	141,3	0,795	0,485	0,828	0,603	0,821	

NEUTRO	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	199	3015	361	489	0,289	0,355	0,893	0,319	0,791		336	2353	1023	352	0,488	0,247	0,697	0,328	0,662	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	132	3209	167	556	0,192	0,441	0,951	0,267	0,822		305	2542	834	383	0,443	0,268	0,753	0,334	0,701	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	56	3373	3	632	0,081	0,949	0,999	0,15	0,844		459	3051	325	229	0,667	0,585	0,904	0,624	0,864	

Crescent Serenade

FICHERO																				
SMOTE	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	507,4	1617	998,7	484,6	0,511	0,337	0,618	0,406	0,589		623,4	1331	1285	368,6	0,628	0,327	0,509	0,43	0,542	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	592,2	1436	1180	399,8	0,597	0,334	0,549	0,428	0,562		658,2	1205	1412	333,8	0,664	0,318	0,46	0,43	0,51	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	633,9	1816	799,9	358,1	0,639	0,442	0,694	0,523	0,679		785,1	1371	1245	206,9	0,791	0,387	0,524	0,52	0,591	

CrescentSerenade																				
UNDER	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	622,4	1283	1333	369,6	0,627	0,318	0,491	0,422	0,528		653,1	1185	1431	338,9	0,658	0,313	0,453	0,425	0,51	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	640	1318	1298	352	0,645	0,33	0,504	0,437	0,543		665,2	1165	1451	326,8	0,671	0,314	0,445	0,428	0,507	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	691,6	1667	948,8	300,4	0,697	0,422	0,637	0,525	0,654		779,3	1412	1204	212,7	0,786	0,393	0,54	0,524	0,607	

NEUTRO	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	332	2035	581	660	0,335	0,364	0,778	0,349	0,656		576	1468	1148	416	0,581	0,334	0,561	0,424	0,567	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	276	2239	377	716	0,278	0,423	0,856	0,336	0,697		557	1524	1092	435	0,561	0,338	0,583	0,422	0,577	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	97	2580	36	895	0,098	0,729	0,986	0,172	0,742		594	1871	745	398	0,599	0,444	0,715	0,51	0,683	

ForTwo

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	283,6	6099	1154	559,4	0,336	0,197	0,841	0,249	0,788	385,7	5143	2110	457,3	0,458	0,155	0,709	0,231	0,683
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	336,4	5866	1387	506,6	0,399	0,195	0,809	0,262	0,766	390,2	5187	2067	452,8	0,463	0,159	0,715	0,237	0,685
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	663,5	6839	414,5	179,5	0,787	0,616	0,943	0,691	0,927	747,4	6413	840,5	95,6	0,887	0,471	0,884	0,615	0,884

ForTwo																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	419,2	5206	2047	423,8	0,497	0,17	0,718	0,253	0,695	420,4	4778	2475	422,6	0,499	0,145	0,659	0,225	0,642
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	391,2	5627	1626	451,8	0,464	0,194	0,776	0,274	0,743	406,5	5144	2109	436,5	0,482	0,162	0,709	0,242	0,684
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	643,7	6790	462,7	199,3	0,764	0,582	0,936	0,66	0,918	743,2	6492	761,5	99,8	0,882	0,494	0,895	0,633	0,894

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	188	6632	621	655	0,223	0,232	0,914	0,228	0,842	358	5439	1814	485	0,425	0,165	0,75	0,237	0,716
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	121	7036	217	722	0,144	0,358	0,97	0,205	0,884	321	5969	1284	522	0,381	0,2	0,823	0,262	0,777
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	104	7246	7	739	0,123	0,937	0,999	0,218	0,908	696	6781	472	147	0,826	0,596	0,935	0,692	0,924

Jive

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	306,4	2907	604,7	549,6	0,358	0,336	0,828	0,347	0,736	416,9	2237	1275	439,1	0,487	0,246	0,637	0,327	0,608
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	361,9	2793	718,9	494,1	0,423	0,335	0,795	0,374	0,722	461,9	2124	1388	394,1	0,54	0,25	0,605	0,341	0,592
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	558,9	2679	833,3	297,1	0,653	0,401	0,763	0,497	0,741	722,6	1768	1744	133,4	0,844	0,293	0,503	0,435	0,57

Jive																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	424,9	2455	1057	431,1	0,496	0,287	0,699	0,363	0,659	454,3	1973	1539	401,7	0,531	0,228	0,562	0,319	0,556
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	404,7	2589	923,4	451,3	0,473	0,305	0,737	0,371	0,685	466,2	2092	1420	389,8	0,545	0,247	0,596	0,34	0,586
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	578,7	2640	872,3	277,3	0,676	0,399	0,752	0,502	0,737	710,3	1842	1670	145,7	0,83	0,298	0,524	0,439	0,584

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	204	3205	307	652	0,238	0,399	0,913	0,298	0,78	380	2422	1090	476	0,444	0,259	0,69	0,327	0,641
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	144	3359	153	712	0,168	0,485	0,956	0,25	0,802	364	2617	895	492	0,425	0,289	0,745	0,344	0,682
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	210	3419	93	646	0,245	0,693	0,974	0,362	0,831	568	2443	1069	288	0,664	0,347	0,696	0,456	0,685

LoungeAway

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	314,7	2151	573,3	385,3	0,45	0,354	0,79	0,396	0,72	399,4	1659	1065	300,6	0,571	0,273	0,609	0,369	0,601
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	368,7	2022	701,6	331,3	0,527	0,344	0,742	0,417	0,698	435,3	1629	1095	264,7	0,622	0,285	0,598	0,39	0,603
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	461,4	2057	667	238,6	0,659	0,409	0,755	0,505	0,736	574	1617	1108	126	0,82	0,341	0,593	0,482	0,68

LoungeAway																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	415,5	1782	941,9	284,5	0,594	0,306	0,654	0,404	0,642	418,4	1550	1174	281,6	0,598	0,263	0,569	0,365	0,575
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	392,1	1867	857,3	307,9	0,56	0,314	0,685	0,402	0,66	439,2	1600	1124	260,8	0,627	0,281	0,587	0,388	0,596
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	492,6	2011	712,9	207,4	0,704	0,409	0,738	0,517	0,731	552,4	1682	1042	147,6	0,789	0,347	0,618	0,482	0,653

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	230	2392	332	470	0,329	0,409	0,878	0,365	0,766	369	1780	944	331	0,527	0,281	0,653	0,367	0,628
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	188	2522	202	512	0,269	0,482	0,926	0,345	0,791	384	1921	803	316	0,549	0,324	0,705	0,407	0,673
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	159	2678	46	541	0,227	0,776	0,983	0,351	0,829	467	1941	783	233	0,667	0,374	0,713	0,479	0,703

RM-J001-MID

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	614,3	2851	666,1	244,7	0,715	0,48	0,811	0,574	0,792	711,6	2114	1403	147,4	0,828	0,336	0,601	0,479	0,646
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	680,8	2675	841,9	178,2	0,793	0,447	0,761	0,572	0,767	734,8	2087	1430	124,2	0,855	0,339	0,593	0,486	0,645
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	661,3	2948	569	197,7	0,77	0,538	0,838	0,633	0,825	695,7	2749	768	163,3	0,81	0,475	0,782	0,599	0,787

RM-J001-MID																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	681,2	2353	1164	177,8	0,793	0,369	0,669	0,504	0,693	718,8	1891	1626	140,2	0,837	0,307	0,538	0,449	0,596
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	717,8	2447	1070	141,2	0,836	0,401	0,696	0,542	0,723	737,6	1995	1522	121,4	0,859	0,326	0,567	0,473	0,624
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	686,7	2910	607,1	172,3	0,799	0,531	0,827	0,638	0,822	696,1	2746	771,4	162,9	0,81	0,474	0,781	0,598	0,788

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	541	3092	425	318	0,63	0,56	0,879	0,593	0,83	693	2252	1265	166	0,807	0,354	0,64	0,492	0,673
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	558	3230	287	301	0,65	0,66	0,918	0,655	0,866	704	2487	1030	155	0,82	0,406	0,707	0,543	0,725
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	219	3489	28	640	0,255	0,887	0,992	0,396	0,847	633	2888	629	226	0,737	0,502	0,821	0,597	0,805

RM-J002-MID

FICHERO																				
SMOTE	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	629,7	5722	1539	237,3	0,726	0,29	0,788	0,415	0,781		732,7	4657	2604	134,3	0,845	0,22	0,641	0,349	0,663	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	695,4	5388	1874	171,6	0,802	0,271	0,742	0,405	0,748		749,2	4598	2663	117,8	0,864	0,22	0,633	0,35	0,658	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	719,6	6461	800,2	147,4	0,83	0,473	0,89	0,603	0,883		732,9	6320	941	134,1	0,845	0,438	0,87	0,577	0,868	

RM-J002-MID																				
UNDER	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	697,1	4837	2424	169,9	0,804	0,223	0,666	0,35	0,681		733,6	4437	2824	133,4	0,846	0,206	0,611	0,332	0,636	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	729,6	5077	2184	137,4	0,842	0,25	0,699	0,386	0,714		742,9	4635	2626	124,1	0,857	0,221	0,638	0,351	0,662	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	759,8	6315	945,8	107,2	0,876	0,446	0,87	0,591	0,87		732,4	6326	934,6	134,6	0,845	0,439	0,871	0,578	0,868	

NEUTRO	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	564	6359	902	303	0,651	0,385	0,876	0,483	0,852		719	4940	2321	148	0,829	0,237	0,68	0,368	0,696	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	572	6734	527	295	0,66	0,52	0,927	0,582	0,899		726	5272	1989	141	0,837	0,267	0,726	0,405	0,738	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	205	7228	33	662	0,236	0,861	0,995	0,371	0,914		679	6518	743	188	0,783	0,477	0,898	0,593	0,885	

RM-J003-MID

FICHERO																				
SMOTE	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	490	2306	523,1	201	0,709	0,484	0,815	0,575	0,794		569,8	1784	1045	121,2	0,825	0,353	0,631	0,494	0,669	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	528,6	2198	631,3	162,4	0,765	0,456	0,777	0,571	0,775		576,9	1782	1047	114,1	0,835	0,355	0,63	0,499	0,67	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	478	2352	477,1	213	0,692	0,5	0,831	0,581	0,804		520,5	2193	635,6	170,5	0,753	0,45	0,775	0,564	0,771	

RM-J003-MID																				
UNDER	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	545,9	1992	836,9	145,1	0,79	0,395	0,704	0,526	0,721		580,7	1693	1136	110,3	0,84	0,338	0,599	0,482	0,646	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	561,9	2057	771,9	129,1	0,813	0,421	0,727	0,555	0,744		584,1	1760	1069	106,9	0,845	0,353	0,622	0,498	0,666	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	507,3	2318	511,2	183,7	0,734	0,498	0,819	0,594	0,803		513,8	2188	640,8	177,2	0,744	0,445	0,773	0,557	0,768	

NEUTRO	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	435	2473	356	256	0,63	0,55	0,874	0,587	0,826		562	1901	928	129	0,813	0,377	0,672	0,515	0,7	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	435	2601	228	256	0,63	0,656	0,919	0,643	0,863		565	2056	773	126	0,818	0,422	0,727	0,557	0,743	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	176	2796	33	515	0,255	0,842	0,988	0,391	0,844		447	2322	507	244	0,647	0,469	0,821	0,543	0,785	

RM-J004-MID

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	638,3	1780	735,3	430,7	0,597	0,465	0,708	0,523	0,675	787,4	1488	1027	281,6	0,737	0,434	0,592	0,546	0,635
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	735,4	1633	881,6	333,6	0,688	0,455	0,649	0,548	0,661	815,9	1400	1115	253,1	0,763	0,423	0,557	0,544	0,618
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	648,4	1734	781,1	420,6	0,607	0,454	0,689	0,519	0,665	591,6	1676	838,8	477,4	0,553	0,414	0,666	0,473	0,635

RM-J004-MID																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	769,6	1433	1082	299,4	0,72	0,416	0,57	0,527	0,614	829,1	1294	1222	239,9	0,776	0,404	0,514	0,532	0,592
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	797,9	1415	1100	271,1	0,746	0,42	0,563	0,538	0,617	830,8	1335	1180	238,2	0,777	0,413	0,531	0,54	0,604
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	708,5	1628	886,8	360,5	0,663	0,444	0,647	0,532	0,652	595,2	1693	821,9	473,8	0,557	0,42	0,673	0,479	0,631

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	500	2098	417	569	0,468	0,545	0,834	0,504	0,725	746	1638	877	323	0,698	0,46	0,651	0,554	0,665
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	451	2225	290	618	0,422	0,609	0,885	0,498	0,747	729	1728	787	340	0,682	0,481	0,687	0,564	0,686
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	34	2505	10	1035	0,032	0,773	0,996	0,061	0,708	417	1968	547	652	0,39	0,433	0,783	0,41	0,665

RM-J005-MID

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	401	2644	630,7	308	0,566	0,389	0,807	0,461	0,764	517,1	2112	1163	191,9	0,729	0,308	0,645	0,433	0,665
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	456,7	2498	777,4	252,3	0,644	0,37	0,763	0,47	0,742	535,1	2088	1187	173,9	0,755	0,311	0,638	0,44	0,655
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	450,2	2871	403,7	258,8	0,635	0,527	0,877	0,576	0,834	477,7	2708	566,7	231,3	0,674	0,457	0,827	0,545	0,8

RM-J005-MID																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	489,2	2172	1103	219,8	0,69	0,307	0,663	0,425	0,668	533,5	1956	1319	175,5	0,752	0,288	0,597	0,417	0,625
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	499,6	2286	989,3	209,4	0,705	0,336	0,698	0,455	0,699	537,9	2017	1259	171,1	0,759	0,299	0,616	0,429	0,641
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	486,7	2811	463,9	222,3	0,686	0,512	0,858	0,587	0,828	470,8	2725	550,4	238,2	0,664	0,461	0,832	0,544	0,801

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	327	2886	389	382	0,461	0,457	0,881	0,459	0,806	495	2228	1047	214	0,698	0,321	0,68	0,44	0,683
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	314	3067	208	395	0,443	0,602	0,936	0,51	0,849	500	2395	880	209	0,705	0,362	0,731	0,479	0,721
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	67	3262	13	642	0,094	0,838	0,996	0,17	0,836	382	2852	423	327	0,539	0,475	0,871	0,505	0,812

RM-J006-MID

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	356,3	2131	610,9	333,7	0,516	0,368	0,777	0,43	0,725	469,1	1607	1135	220,9	0,68	0,292	0,586	0,409	0,605
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	411,5	1979	762,7	278,5	0,596	0,35	0,722	0,441	0,697	510,7	1478	1265	179,3	0,74	0,288	0,539	0,414	0,575
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	394,2	2073	668,7	295,8	0,571	0,371	0,756	0,45	0,719	429,7	1919	822,8	260,3	0,623	0,343	0,7	0,442	0,684

RM-J006-MID																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	459,7	1645	1097	230,3	0,666	0,295	0,6	0,409	0,613	506,7	1389	1353	183,3	0,734	0,272	0,507	0,397	0,552
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	481,4	1606	1136	208,6	0,698	0,298	0,586	0,417	0,608	518,7	1417	1325	171,3	0,752	0,281	0,517	0,409	0,564
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	392,8	2076	665,8	297,2	0,569	0,371	0,757	0,449	0,719	428	1955	787,2	262	0,62	0,352	0,713	0,449	0,694

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	263	2361	381	427	0,381	0,408	0,861	0,394	0,765	425	1793	949	265	0,616	0,309	0,654	0,412	0,646
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	245	2478	264	445	0,355	0,481	0,904	0,409	0,793	422	1909	833	268	0,612	0,336	0,696	0,434	0,675
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	172	2539	203	518	0,249	0,459	0,926	0,323	0,79	361	2112	630	329	0,523	0,364	0,77	0,43	0,721

RM-J007-MID

FICHERO																		
SMOTE	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	354,4	3086	768,2	463,6	0,433	0,316	0,801	0,365	0,736	467	2578	1276	351	0,571	0,268	0,669	0,365	0,652
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	425,4	2957	896,7	392,6	0,52	0,322	0,767	0,398	0,724	508,3	2461	1393	309,7	0,621	0,267	0,638	0,374	0,635
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	414,1	3425	429,4	403,9	0,506	0,491	0,889	0,498	0,822	458,6	3281	573,5	359,4	0,561	0,444	0,851	0,496	0,8

RM-J007-MID																		
UNDER	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	481,9	2546	1308	336,1	0,589	0,269	0,661	0,37	0,648	511,3	2304	1551	306,7	0,625	0,248	0,598	0,355	0,602
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	489,2	2682	1172	328,8	0,598	0,294	0,696	0,395	0,679	514,2	2436	1418	303,8	0,629	0,266	0,632	0,374	0,631
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
MEDIA	382,7	3487	366,9	435,3	0,468	0,511	0,905	0,488	0,828	445,1	3273	580,8	372,9	0,544	0,434	0,849	0,483	0,796

NEUTRO	ESTANDAR									EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	232	3440	414	586	0,284	0,359	0,893	0,317	0,786	417	2812	1042	401	0,51	0,286	0,73	0,366	0,691
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	189	3696	158	629	0,231	0,545	0,959	0,324	0,832	379	3106	748	439	0,463	0,336	0,806	0,39	0,746
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu
PRUEBA	128	3765	89	690	0,156	0,59	0,977	0,247	0,833	388	3366	488	430	0,474	0,443	0,873	0,458	0,804

RM-J008-MID

FICHERO																				
SMOTE	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	261,2	3359	795,4	328,8	0,443	0,247	0,809	0,317	0,763		321,5	2699	1455	268,5	0,545	0,181	0,65	0,272	0,637	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	302,4	3188	965,8	287,6	0,513	0,238	0,768	0,325	0,736		349,9	2670	1484	240,1	0,593	0,191	0,643	0,289	0,637	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	280	3724	430,1	310	0,475	0,394	0,896	0,431	0,844		314,3	3565	589,4	275,7	0,533	0,348	0,858	0,421	0,818	

RM-J008-MID																				
UNDER	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	343,5	2805	1349	246,5	0,582	0,203	0,675	0,301	0,664		348,2	2450	1704	241,8	0,59	0,17	0,59	0,264	0,59	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	346,3	3028	1127	243,7	0,587	0,235	0,729	0,336	0,711		352	2685	1469	238	0,597	0,193	0,646	0,292	0,64	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	269,7	3754	399,7	320,3	0,457	0,403	0,904	0,428	0,848		293,4	3573	580,8	296,6	0,497	0,336	0,86	0,401	0,815	

NEUTRO	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	175	3693	461	415	0,297	0,275	0,889	0,285	0,815		282	2954	1200	308	0,478	0,19	0,711	0,272	0,682	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	138	3943	211	452	0,234	0,395	0,949	0,294	0,86		262	3348	806	328	0,444	0,245	0,806	0,316	0,761	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	83	4081	73	507	0,141	0,532	0,982	0,223	0,878		249	3713	441	341	0,422	0,361	0,894	0,389	0,835	

RM-J009-MID

FICHERO																				
SMOTE	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	352	3208	956,9	243	0,592	0,269	0,77	0,37	0,748		417	2670	1495	178	0,701	0,218	0,641	0,333	0,648	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	392,1	2988	1177	202,9	0,659	0,25	0,717	0,362	0,71		446,9	2477	1688	148,1	0,751	0,209	0,595	0,327	0,614	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	382,9	3537	628,2	212,1	0,644	0,379	0,849	0,477	0,823		419	3360	804,7	176	0,704	0,342	0,807	0,461	0,794	

RM-J009-MID																				
UNDER	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	420	2612	1553	175	0,706	0,213	0,627	0,327	0,637		446,3	2377	1788	148,7	0,75	0,2	0,571	0,316	0,593	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	441,1	2686	1479	153,9	0,741	0,23	0,645	0,351	0,657		452,6	2386	1779	142,4	0,761	0,203	0,573	0,32	0,596	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	382,3	3585	579,6	212,7	0,643	0,397	0,861	0,491	0,834		404,5	3396	769,3	190,5	0,68	0,345	0,815	0,457	0,796	

NEUTRO	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	267	3572	593	328	0,449	0,31	0,858	0,367	0,807		391	2898	1267	204	0,657	0,236	0,696	0,347	0,691	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	221	3806	359	374	0,371	0,381	0,914	0,376	0,846		371	3039	1126	224	0,624	0,248	0,73	0,355	0,716	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	189	4009	156	406	0,318	0,548	0,963	0,402	0,882		249	3713	441	341	0,422	0,361	0,894	0,389	0,835	

RM-J010-MID

FICHERO																				
SMOTE	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	278,2	3004	952,8	292,8	0,487	0,226	0,759	0,309	0,725		329,3	2556	1401	241,7	0,577	0,19	0,646	0,286	0,637	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	320,7	2825	1132	250,3	0,562	0,221	0,714	0,317	0,695		353,9	2432	1525	217,1	0,62	0,188	0,615	0,289	0,615	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	294	3465	491,7	277	0,515	0,374	0,876	0,433	0,83		350,9	3303	653,7	220,1	0,615	0,349	0,835	0,445	0,807	

RM-J010-MID																				
UNDER	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	349,7	2456	1501	221,3	0,612	0,189	0,621	0,289	0,62		361,6	2325	1632	209,4	0,633	0,181	0,588	0,282	0,593	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	362,1	2596	1361	208,9	0,634	0,21	0,656	0,316	0,653		353	2483	1474	218	0,618	0,193	0,628	0,294	0,626	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA	305,2	3465	492,1	265,8	0,535	0,383	0,876	0,446	0,833		328,5	3336	620,7	242,5	0,575	0,346	0,843	0,432	0,805	

NEUTRO	ESTANDAR										EXTENDIDO									
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	183	3439	518	388	0,32	0,261	0,869	0,288	0,8		299	2828	1129	272	0,524	0,209	0,715	0,299	0,691	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	147	3675	282	424	0,257	0,343	0,929	0,294	0,844		263	3090	867	308	0,461	0,233	0,781	0,309	0,741	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
PRUEBA	59	3879	78	512	0,103	0,431	0,98	0,167	0,87		267	3492	465	304	0,468	0,365	0,882	0,41	0,83	

Tabla 3: Resultados individuales por fichero de audio.

Una vez presentados los resultados individuales es momento de realizar un primer análisis de los mismos, siempre intentando tener en cuenta los distintos puntos de vista con los que se puede enfocar el problema. Ha habido una combinación de tipo de etiquetado y clasificador cuyos resultados han destacado con respecto a los producidos por cualquier otra combinación. Se trata de la técnica de etiquetado estándar y las SVM. Se han obtenido resultados dispares en función de los distintos conjuntos de entrenamiento afectados por SMOTE, *Undersampling* y Neutro utilizados para entrenar al clasificador. La combinación formada por la técnica de etiquetado estándar, la estrategia de re-muestreo neutra y el clasificador SVM ha resultado ser, generalmente, la que ha producido valores más altos de precisión de toda la experimentación realizada además de obtener buenos resultados en cuanto a *accuracy*. El inconveniente principal que presentan las soluciones derivadas de esta combinación es que producen resultados muy sesgados hacia la clase negativa (no *onset*), lo cual se traduce en que se dejan por el camino muchísimos *frames* que deberían ser detectados como *onset*. Dicho de otra forma, logra muy buenos resultados en la clase negativa y en aquellos indicadores que dependen de la tasa global de aciertos, pero sus resultados en la clase positiva (detección de onsets), que es la que realmente interesa, son muy bajos. Estos resultados confirman que medidas como el *accuracy*, analizadas fuera de contexto, pueden conducir a resultados engañosos. Las técnicas de re-muestreo SMOTE y *Undersampling* aplicadas sobre los conjuntos de entrenamiento han corregido esta tendencia, aumentando los índices de TP pero a cambio de reducir un poco la precisión en la detección, ya que también han incrementado los índices de FP. Generalmente, son preferibles resultados más equilibrados como los generados por estas técnicas, obteniendo rendimientos sensiblemente mejores en la clase positiva, a costa de causar un perjuicio moderado en la negativa.

En las Figuras 7, 8 y 9 se resume gráficamente los resultados producidos por las combinaciones de técnica de etiquetado estándar y clasificador SVM en función de las tres estrategias de re-muestreo para las medidas *recall*, precisión y *accuracy*.

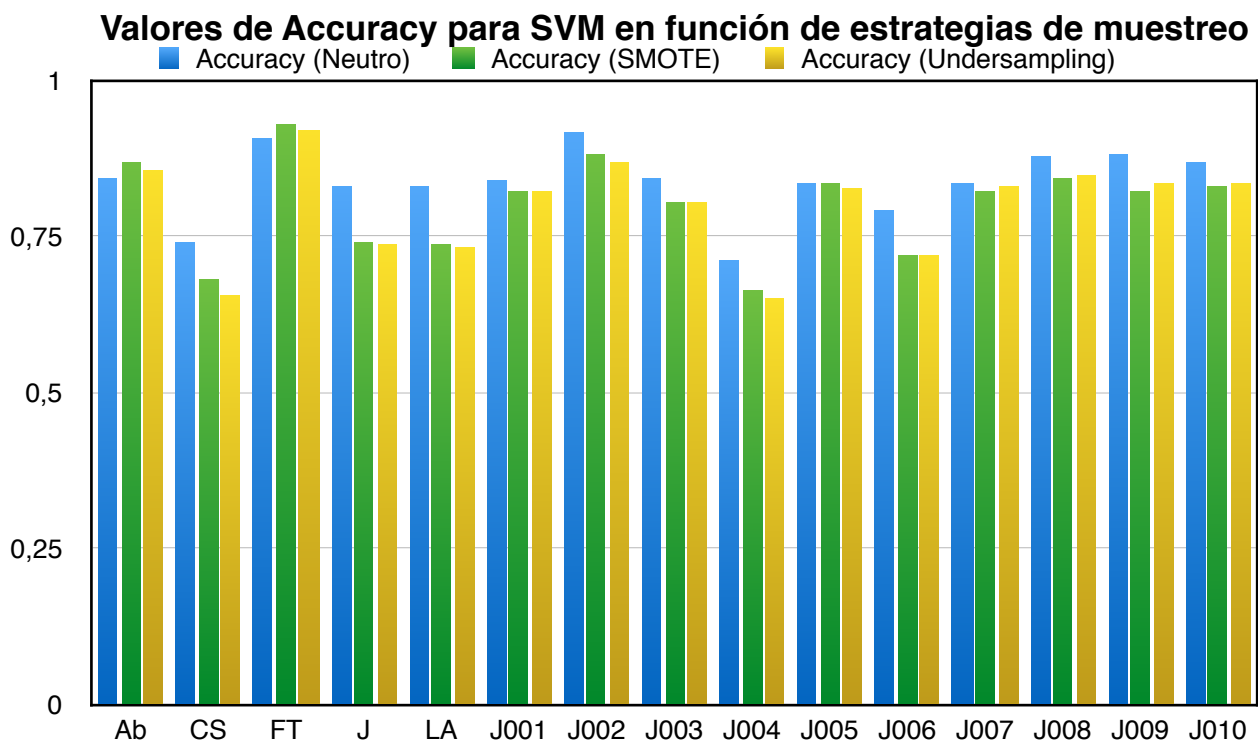


Figura 7: Valores de *Accuracy* para cada fichero de audio, producidos por la combinación de la técnica de etiquetado estándar y el clasificador SVM en función de las distintas estrategias de re-muestreo utilizadas.

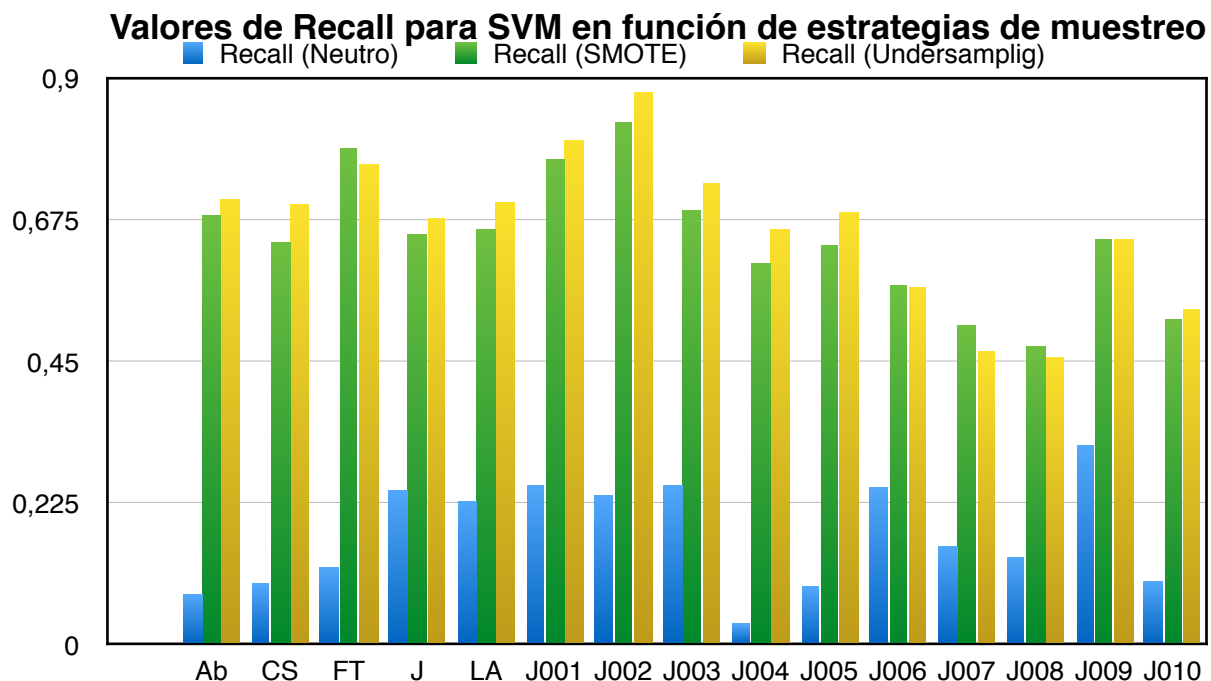


Figura 8: Valores de *Recall* para cada fichero de audio, producidos por la combinación de la técnica de etiquetado estándar y el clasificador SVM en función de las distintas estrategias de re-muestreo utilizadas.

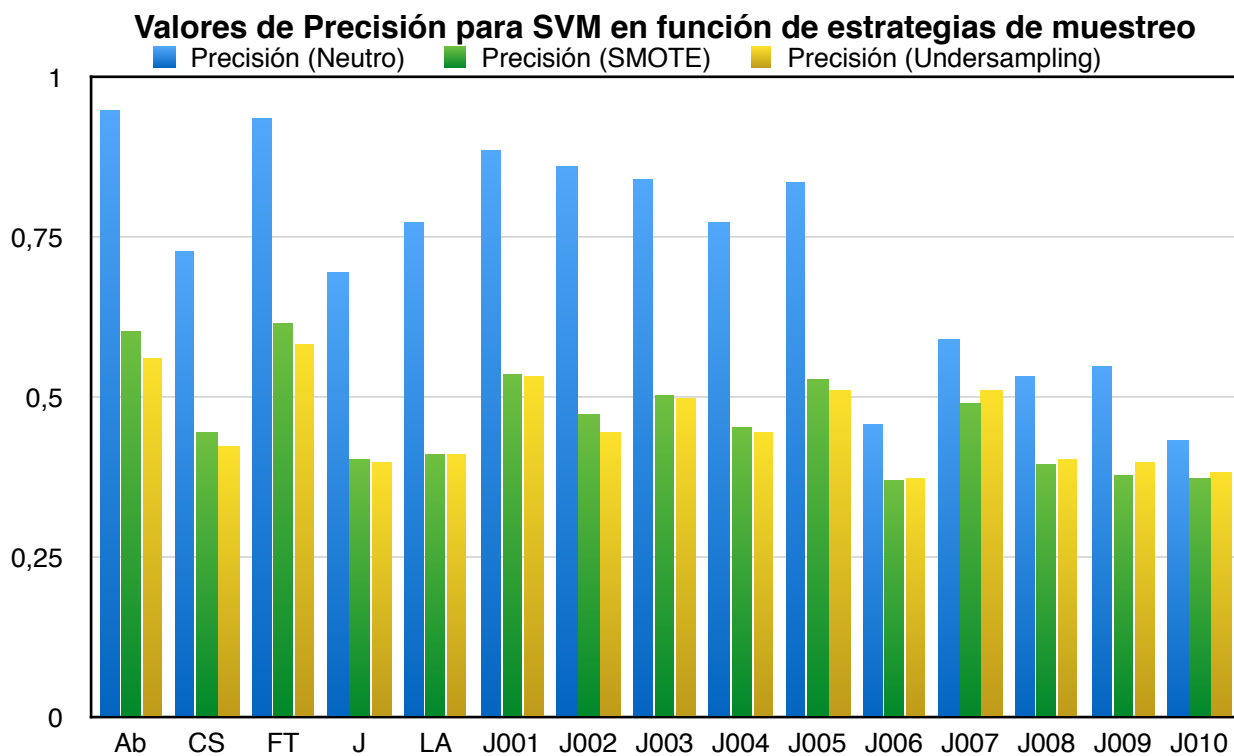


Figura 9: Valores de Precisión para cada fichero de audio, producidos por la combinación de la técnica de etiquetado estándar y el clasificador SVM en función de las distintas estrategias de re-muestreo utilizadas.

En la Figura 7 se observan los valores de *accuracy* obtenidos para las distintas estrategias de re-muestreo. Como tendencia general, los conjuntos de entrenamiento neutros han obtenido valores mayores, aunque hay varias excepciones como son los casos de Abyss y ForTwo donde SMOTE y *Undersampling* han sido mejores. Más adelante se ofrecen posibles razones que explicarían este comportamiento.

Por un lado, en la Figura 8 se puede observar un incremento muy considerable del *Recall* al entrenar el clasificador SVM con conjuntos de entrenamiento afectados por estrategias de re-muestreo como SMOTE y *Undersampling*. Una conclusión interesante que se extrae es que los métodos de re-muestreo mejoran mucho el rendimiento de los detectores pues, manteniendo prácticamente el mismo *accuracy* que el método neutro, mejoran muchísimo el *recall* (que es la tasa de aciertos en la clase de interés, es decir, la tasa de detección de *onsets*).

Por otro lado, en la Figura 9 se puede comprobar, de manera gráfica, que los conjuntos de entrenamiento neutros generan resultados globalmente más precisos que los afectados por SMOTE y *Undersampling*, aunque más sesgados hacia la clase negativa.

Finalmente destacar el archivo de audio *ForTwo*, pues ha dado lugar a los mejores resultados de detección automática de *onsets* de todo el trabajo gracias a las combinación <Estándar-SMOTE-SVM>. Una posible razón podría encontrarse en la relación que se establece entre el número de *frames* y el número de *onsets* anotados manualmente por el oyente experto (ver Tabla 1), por lo que deducimos que se trata de un audio con un *tempo* largo, donde el espacio de tiempo entre cada nota es mayor que en el resto de audios y puede existir menor incertidumbre a la hora de anotar o de detectar el inicio de cada nota. Dada la naturaleza secuencial de los *frames*, un desbalance en el conjunto de pruebas

entre las clases del problema, podría estar relacionado con un mayor espaciado de los inicios de notas y, por tanto, con una mayor probabilidad de detectarlos. Otro posible motivo puede ser la misma sencillez de esta obra. En primer lugar, se trata de interpretaciones realizadas mediante un único instrumento como es el piano, de manera que, físicamente, el problema de detección se delimita a un sonido y unas frecuencias singulares. En segundo lugar, existe una escasa percepción de distintas notas sonando al mismo tiempo, siendo la mayoría de ellas de duración breve y producidas como acompañamiento melódico, por lo que las transiciones entre notas son más fáciles de detectar. Estos hechos facilitan la compactación de la estructura de la clase *onset* y, de esta forma, que SMOTE pueda tener más éxito, ya que esta estrategia de re-muestreo crea nuevas muestras a partir de las ya existentes y, si éstas no son ruidosas, las nuevas muestras representarán con mayor fidelidad las peculiaridades físicas de las originales.

4.2. Resultados generales

Se han resumido las quince tablas de resultados individuales en una única tabla resumen. Este nuevo bloque muestra una visión global de todo el proceso de clasificación producido. Para organizar los nuevos resultados se ha utilizado el mismo esquema usado para estructurar los resultados de cada fichero individual (Figura 6). Cada fila de datos en esta nueva tabla se obtiene calculando la media aritmética de los datos de la misma fila en cada una de las quince tablas individuales. Como se puede observar, se ignoran los índices básicos (TP, TN, FP, FN) y se tienen en cuenta únicamente las medidas de rendimiento (*recall*, precisión, especificidad, *f-measure* y *accuracy*). A continuación se muestran en la Tabla 4 los resultados globales de todo el proceso de experimentación realizado en este trabajo.

Resultados generales																			
FICHERO																			
SMOTE	ESTANDAR										EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA					0,53	0,34	0,78	0,4	0,74					0,65	0,27	0,63	0,38	0,63	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA					0,6	0,33	0,73	0,42	0,71					0,68	0,27	0,6	0,38	0,63	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA					0,64	0,47	0,83	0,54	0,8					0,72	0,4	0,75	0,51	0,73	

Resultados generales																			
UNDER	ESTANDAR										EXTENDIDO								
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA					0,65	0,28	0,65	0,39	0,65					0,68	0,25	0,57	0,37	0,54	
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA					0,66	0,3	0,67	0,41	0,67					0,66	0,3	0,67	0,41	0,63	
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu	
MEDIA					0,67	0,46	0,82	0,54	0,8					0,67	0,46	0,82	0,54	0,8	

NEUTRO																					
ESTANDAR											EXTENDIDO										
KNN1	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		
PRUEBA					0,4	0,39	0,87	0,39	0,79						0,61	0,28	0,68	0,38	0,67		
KNN3	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		
PRUEBA					0,35	0,49	0,93	0,4	0,83						0,59	0,32	0,73	0,41	0,71		
SVM	TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		TP	TN	FP	FN	Recall	Preci	Speci	f_me	Accu		
PRUEBA					0,17	0,72	0,98	0,27	0,84						0,6	0,44	0,83	0,5	0,75		

Tabla 4: Resultados generales de la experimentación realizada.

Por lo que respecta al análisis general, en primer lugar, como se ha podido deducir de los resultados individuales, el algoritmo de detección SVM ha cosechado siempre los mejores resultados en cuanto a precisión en la detección de *onsets*, especificidad y exactitud (*accuracy*) de la clasificación, y en todas las medidas de rendimiento (incluyendo *recall* y *f-measure*) en combinaciones distintas a la <Estándar-Neutro>. Para esta combinación particular, el algoritmo K-NN para sus versiones K=1,3 ha obtenido mejores resultados únicamente para *recall* y *f-measure* que las SVM, por lo que se considera que ha ofrecido un rendimiento menos sesgado que las máquinas de vector soporte.

Por otro lado, desde el punto de vista de las técnicas de etiquetado, destacar que las versiones extendidas han conseguido un aumento significativo del número de TP, pero a cambio de incrementar desproporcionadamente el número de FP, es decir, se etiqueta muchos más *frames* como *onset* cuando realmente no lo son, incluso este comportamiento se ve incrementado al entrenar al clasificador con conjuntos de entrenamientos afectados por estrategias de re-muestreo como SMOTE y *Undersampling*. De esta forma el tipo de etiquetado estándar ofrece mejores resultados que el etiquetado extendido en cuanto a medidas como precisión, especificidad, *f-measure* y exactitud en la clasificación.

Para acabar con el análisis general y desde la perspectiva de las estrategias de re-muestreo, recalcar lo ya mencionado en la discusión de la sección anterior. Los conjuntos de entrenamiento formados por ficheros de audio originales (neutros) han ayudado a construir modelos de clasificación con mayor valor de precisión, aunque más sesgados hacia la clase negativa (valores de *recall* muy pobres). Generan muy pocos FP pero también muy pocos TP (*onsets*). Las estrategias SMOTE y *Undersampling* mejoran este déficit (*recall* y *f-measure*) a cambio de un descenso de la precisión. Finalmente, en la Figura 10 se observa una comparativa de los resultados generales, donde se puede apreciar gráficamente las tendencias comentadas con anterioridad.

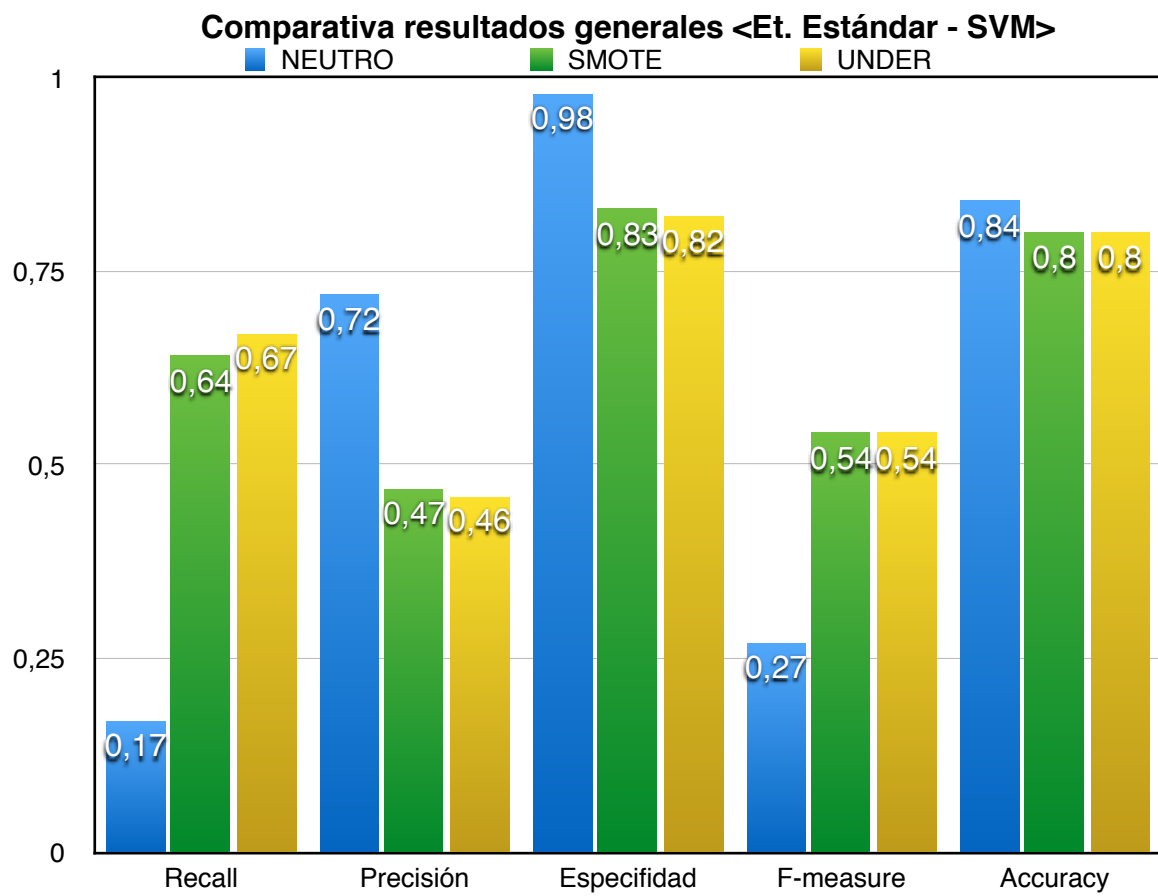


Figura 10: Comparativa de los resultados generales producidos por la combinación de la técnica de etiquetado estándar y el clasificador SVM en función de las distintas estrategias de re-muestreo utilizadas.

Capítulo 5: Conclusiones y trabajo futuro

La detección automática de *onsets* es una poderosa herramienta en el campo de la informática musical y tiene, como se ha explicado a lo largo de este trabajo, un gran número de campos de aplicación. A día de hoy la detección automática de *onsets* sigue siendo un problema abierto. Existen multitud de enfoques al problema desde hace muchos años, pero a día de hoy, aunque se ha mejorado bastante en porcentajes de detección correcta gracias a la potencia de los ordenadores, aún no existe una metodología de referencia. Además, la naturaleza y características de las señales musicales es muy amplia y diferente, y no todos los detectores funcionan bien en todos los casos.

Este estudio forma parte de un trabajo de investigación que se ha llevado a cabo, prácticamente, a lo largo de todo un año. Corresponde con la continuación del Trabajo de Iniciación a la Investigación (TII) donde, a grandes rasgos, se realizó una revisión del estado del arte más reciente y se definió y se caracterizó el problema. Además, fue en este TII donde se introdujo un estudio de soluciones al problema de detección automática de *onsets* que ha sido llevado a cabo en este TFM.

El principal objetivo ha sido desarrollar distintos escenarios de experimentación para comprobar como las técnicas de aprendizaje automático pueden ayudar a resolver este problema.

En primer lugar, se ha producido una transformación de los datos mediante distintas estrategias de re-muestreo sobre el conjunto de ficheros de audio que se han usado como conjunto de entrenamiento, con el objetivo de corregir la situación inicial de desequilibrio entre clases. En segundo lugar, se ha seleccionado un conjunto de algoritmos de clasificación, bien conocidos en el ámbito del aprendizaje automático, para iniciar un proceso de detección automática de *onsets*. Éstos han sido el algoritmo de los K vecinos más cercanos (K-NN) y las máquinas de vectores soporte (SVM), cuyo aprendizaje se ha realizado sobre conjuntos de entrenamiento creados a partir de distintas técnicas de etiquetado y estrategias de re-muestreo, con el objetivo de, en primer lugar, analizar el rendimiento de cada clasificador y, finalmente, encontrar la combinación que produzca la mejor detección posible.

De este trabajo se puede concluir que las SVM son un buen algoritmo de aprendizaje y detección para este problema, obteniendo índices más que aceptables de TP, TN, FP y FN. Por otro lado, la técnica de etiquetado extendida, ideada en el pasado TII como alternativa a la visión más estricta de alineación de *frames*, no ha dado los resultados que se esperaban. Asumir cierta incertidumbre en la elección, por parte del oyente experto, del instante de tiempo donde presumiblemente se inicia una nota, ha derivado en modelos de clasificación más complejos y menos precisos. De esta forma, se ha ayudado a aumentar el número de *onsets* correctamente etiquetados, pero se ha producido un exceso desproporcionado de muestras mal etiquetadas como *onsets* (FP). Sería interesante estudiar esta técnica de etiqueta extendida usando umbrales más pequeños (<50 ms), de modo que se cubra más moderadamente el principio de incertidumbre en las anotaciones.

Contrariamente, la técnica de etiquetado estándar ha pecado de rigurosidad. Los clasificadores han generado modelos que han considerado muy pocos *frames* como *onset*, (muy pocos TP). De esta forma, se han producido clasificaciones muy sesgadas hacia la clase negativa, pero, afortunadamente, han podido ser corregidas mediante distintas estrategias de re-muestreo como SMOTE o *Undersampling*.

En definitiva, las técnicas de aprendizaje automático pueden dar una buena solución al problema de detección automática de *onsets*. El principal inconveniente es el origen subjetivo sobre el que se cimienta el estudio y es que, a día de hoy, se necesita del conocimiento y la experiencia de un oyente experto para generar el primer modelo de clasificación. De todas formas, los resultados podrían mejorarse mediante un estudio de selección de características y eliminación de información redundante y, lógicamente, mediante la experimentación con otros detectores, técnicas de etiquetado y estrategias de re-muestreo. Además, ya todo como futuro trabajo, sería interesante un estudio para evaluar el grado de coincidencia de distintos expertos en el proceso de obtener anotaciones manuales. El objetivo sería conocer la fiabilidad de las anotaciones manuales y, con esta información, definir un umbral apropiado y así, posiblemente, evitar los excesos que hemos experimentado fijándolo en 50 ms.

Referencias

- [1] Bello, J. P., Daudet, L., Abdallah, S. A., Duxbury, C., Davies, M. E., & Sandler, M. B.: A Tutorial on Onset Detection in Music Signals. *IEEE Transactions on Speech and Audio Processing*, 13 (5), 1035–1047 (2005)
- [2] Juan P. Bello, Chris Duxbury, Mike Davies, and Mark Sandler. On the Use of Phase and Energy for Musical Onset Detection in the Complex Domain. *IEEE Signal Processing Letters*, vol. 11, no. 6, pp. 553-555, 2004.
- [3] Chuan, C.-H., & Chew, E.: Audio Onset Detection Using Machine Learning Techniques: The Effect and Applicability of Key and Tempo Information (Tech. Rep.). California, USA: University of Southern California Computer Science Department (2008).
- [4] F. Gouyon, G. Widmer, X. Serra, and A. Flexer. Acoustic cues to beat induction: a machine learning perspective, *Music Perception*, vol. 24, issue 2, pp. 177-188, 2006.
- [5] R. Dannenberg. Bootstrap learning for accurate onset detection, *Machine Learning*, vol. 65, no. 2-3, pp. 457-471, 2006.
- [6] Böck, Sebastian, Florian Krebs, and Markus Schedl. Evaluating the Online Capabilities of Onset Detection Methods. *ISMIR*. 2012.
- [7] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall and W. Philip Kegelmeyer, SMOTE: Synthetic Minority Over-sampling Technique, *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002.
- [8] Liu, Alexander Yun-chung. The effect of oversampling and undersampling on classifying imbalanced text datasets. Diss. The University of Texas at Austin, 2004.
- [9] Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, 4(2), 1883.
- [10] Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters*, 9(3), 293-300.